Some polynomial extremal problems which emerged in the twentieth century

Bahman Saffari

Université de Paris-Sud
Bahman.Saffari@math.u-psud.fr

ABSTRACT. Most of the "extremal problems" of Harmonic (or Fourier) Analysis which emerged before the year 2000 were actually born in the twentieth century, and their emergences were scattered throughout that century, including the two world war periods. A great many of these problems pertain to polynomials, trigonometric polynomials and (finite) exponential sums. Writing a reasonably complete monograph on this huge subject (even if we choose to restrict it to polynomials only) would be a monumental task, although the literature does indeed contain some valuable monographs on various aspects of the subject. The present text just *touches upon* a number of extremal problems on polynomials and trigonometric polynomials, with the hope of expanding this same text in the near future to a much larger version, and ultimately to a "reasonably complete" monograph (but only with the help of other mathematicians.)

The theory of polynomials on the unit circle is, of course, part of classical Fourier Analysis, studied with the tools of real and complex analysis. But it also leads to studying polynomials on the (cyclic) finite subgroups of the unit circle, and this is part of Fourier Analysis on finite groups. In many ways this leads to *cyclotomy*, which is part of Number Theory and Algebra. Also, some *combinatorial designs* (cyclic difference sets) show up in connection with this study. Thus the analysis of polynomials and trigonometric polynomials, even in one single variable, is at the crossroad of many important areas of contemporary mathematics. It is also much connected with some areas of engineering, such as signal processing.

1. Introduction

Whatever is worth doing is worth doing badly
Gilbert Keith Chesterton

Everybody writes, nobody reads
Paul Erdös

I have borrowed from Z. A. Melzak's excellent book "Companion to Concrete Mathematics" (volume II) [26] the first of the above two epigraphs: "*Whatever is worth doing is worth doing badly*". Indeed this sentence is a most relevant epigraph for the present paper, and *even more so* for any attempt to write a fairly complete monograph on the topics touched upon in this paper: I shall explain this in some detail in Section 1.2 (on the "peculiarities and aims of this paper"). The second epigraph: "*Everybody writes, nobody reads*" is a frightening truth so concisely enunciated by the great Paul Erdös (1913–1996), and seems to be an explanation for a good many of the evils that are infecting current science research, whether pure or applied. More about this later.

1.1. *A little history*

Let me start with brief historical remarks concerning (only the origins of) the subject: "extremal problems on polynomials and trigonometric polynomials: a century of progress", as this was the (much too ambitious) initial title I had given the organizers of this ASI, before toning it down to the present title.

To the best of my knowledge (at the time this paper is being written, *i.e.*, in the first week of November 2000), the subject of "*extremal problems on polynomials and trigonometric polynomials*", or at least the *analytic* theory of this subject, seems to find its main roots in two (somewhat distinct) fertile grounds: on one hand the nineteenth century theory of *trigonometric and Fourier series*, and on the other hand the nineteenth century theory of *approximation and interpolation*. The *discrete* aspects of the subject are much related to number theory and can even be tracked down (somewhat loosely) to the time of Gauss and Lagrange. But the century-old *golden age* of the subject seems to have really started around 1900 (*precisely* in 1889 and in 1911) with two *totally independent* major results: first *Markov's inequality*, proved in 1887 in a very special case by the chemist Mendeleiev [27] and in 1889 by his mathematician friend A. A. Markov [25] in the general case, and then *Bernstein's inequality* (proved in 1911 by S. N. Bernstein [3] *in a somewhat weaker form* than what is presently known as "Bernstein's inequality".)

Markov's inequality (in its modern formulation) says that

$$\|P'\|_{[-1,1]} \leqslant n^2 \|P\|_{[-1,1]} \tag{1.1}$$

whenever $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$ is a polynomial with complex coefficients, $P'(X)$ is its derivative, and

$$\|P\|_{[-1,1]} := \max_{-1 \leqslant x \leqslant 1} |P(x)|. \tag{1.2}$$

(1.1) is an equality if we take $P(X) = T_n(X)$ where the Chebishev (or Tchebyshev) polynomial $T_n(X)$ is defined by $\cos nu = T_n(\cos u)$. Markov's inequality had been proved in the special case $n = 2$ as early as 1887 by the chemist Mendeleiev who studied it in connection with a problem on substances dissolved in a liquid [27]. He then mentioned it to A. A. Markov who became very interested and proved [25] the general case of (1.1) shortly after Mendeleiev's work [27]. For more historical details, see *for example* [31] or [32].

Unlike Markov's inequality (which, as we saw, originated from Mendeleiev's research in chemistry), *Bernstein's inequality* originated from pure mathematics (approximation theory). Indeed, at the beginning of the twentieth century, the Belgian mathematician De la Vallée Poussin [11] asked whether any piecewise linear continuous function defined on a compact interval of $\mathbb{R}$ could be approximated by polynomials of degree n with a (uniform) error $o(1/n)$ as $n \to \infty$. With the less drastic error $O(1/n)$, the (affirmative) answer had been given by De la Vallée Poussin himself. In a celebrated memoir (which was awarded a prize by the Royal Academy of Belgium on 15 December 1911), S. N. Bernstein [3] answered De la Vallée Poussin's question *negatively*: He proved that the best (uniform) approximation of the function $|x|$ on $[-1, 1]$ by a polynomial of degree $2n$, $(n \geqslant 1)$, lies between $\frac{\sqrt{2}-1}{4} \cdot \frac{1}{2n-1}$ and $\frac{2}{\pi} \cdot \frac{1}{2n+1}$. Bernstein's proof of this theorem heavily uses the following inequality which he proves in the same memoir [3]: Whenever $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$,

$$|P'(x)| \leqslant \frac{n}{\sqrt{1-x^2}} \|P\|_{[-1,1]} \quad (-1 < x < 1). \tag{1.3}$$

Actually, Bernstein [3] proved (1.3) only for $P(X) \in \mathbb{R}[X]$, but the extension of (1.3) to $P(X) \in \mathbb{C}[X]$ is straightforward: Indeed, if $P(X) \in \mathbb{C}[X]$, put $Q(x) := \Re(P(x))$ for $-1 \leqslant x \leqslant 1$. Choose $\alpha \in \mathbb{R}$ so that $e^{i\alpha} P'(x)\sqrt{1-x^2}$ attains the maximum $\|P'(x)\sqrt{1-x^2}\|_{[-1,1]}$ at, say, $x = x_0$. Then

$$\begin{aligned} \|P'(x)\sqrt{1-x^2}\|_{[-1,1]} &= e^{i\alpha} P'(x_0)\sqrt{1-x_0^2} \\ &= Q'(x_0)\sqrt{1-x_0^2} \leqslant n\|Q\|_{[-1,1]} \leqslant n\|P\|_{[-1,1]}, \end{aligned}$$

as claimed.

Putting $\|f\|_\infty := \max_{t \in \mathbb{R}} |f(t)|$ where $f(t)$ is a real trigonometric polynomial

$$f(t) = A_0 + \sum_{k=1}^{n} (A_k \cos kt + B_k \sin kt), \tag{1.4}$$

the inequality

$$\|f'\|_\infty \leqslant n\|f\|_\infty \tag{1.5}$$

(where equality holds if and only if $f(t) = A\cos nt + B\sin nt$) is nowadays known as "*Bernstein's inequality*" although Bernstein never proved this result. Let us explain this. The change of variable $x = \cos t$ shows that any real cosine polynomial

$$f(t) = \sum_{k=0}^{n} A_k \cos kt \tag{1.6}$$

can be written in the form $f(t) = P(\cos t)$ with $P(X) \in \mathbb{R}[X]$ as in (1.3), and vice versa. Thus, apart from the equality case, (1.5) is indeed equivalent to (1.3) *in the case when* $f(t)$ *is a cosine polynomial* (1.6). So Bernstein [3] did state and prove (1.5) for all cosine polynomials. Then by a quite complicated (although very interesting) argument he showed that from the truth of (1.5) for all *cosine* polynomials one can deduce the truth of (1.5) for all *sine* polynomials of the form

$$f(t) = \sum_{k=1}^{n} B_k \sin kt. \tag{1.7}$$

Actually, his argument, as presented in [3], had a gap that Bernstein did correct in his famous book [4] published fourteen years later, in 1926.

Now, having proved (1.5) for all cosine and sine polynomials, Bernstein could only conclude that in the general case of "mixed" real-valued trigonometric polynomials (1.4) one has

$$\|f'\|_\infty \leqslant 2n\|f\|_\infty. \tag{1.8}$$

This is not as good as (1.5), and equality in (1.8) *never* holds unless $f(t) \equiv 0$.

Who supplied the first proof of "Bernstein's inequality" (1.5) for "mixed" trigonometric polynomials (1.4)? In his 1926 book [4], Bernstein stated that the first proof was supplied by E. Landaw [20] in a personal letter he sent Bernstein shortly after the publication of Bernstein's original memoir [3]. In the book [4] Bernstein did give Landau's proof, which consists of a simple and elegant argument showing that the truth of (1.5) for all *sine* polynomials (1.7) in fact implies the truth of (1.5) for all "mixed" polynomials (1.4). Thus Bernstein had done most of the hard work by proving (1.5) for all *sine* polynomials, but had missed Landau's simple argument leading to the general case! (Such things happen quite often.)

So Landau's 1912 proof of "Bernstein's inequality" (1.5) remained unpublished until 1926, when Bernstein's book [4] appeared. However, in a 1914 paper devoted to conjugate trigonometric series, Fejér [15] states the truth of Bernstein's inequality (1.5) in the general case, and gives a very important application of it. Fejér's proof of Bernstein's inequality appeared two years later, in 1916, in a paper of Fekete [16] who indeed does attribute the proof to Fejér. It is thus plausible that the first proofs of "Bernstein's inequality" (1.5) for all real

"mixed" trigonometric polynomials (1.4) were found *independently* by Landau [20] and by Fejér [15], at about the same time, and that both of them communicated their proofs privately to fellow mathematicians (Bernstein and Fekete, respectively.) However, note that in the same year 1914 when Fejér's paper [15] appeared, the brothers M. Riesz [35] and F. Riesz Riesz, F. [34] published two new proofs of Bernstein's inequality (1.5), entirely different from each other and entirely different from those of Landau [20] and of Fejér [15].

Nowadays there are numerous generalizations, extensions and refinements of (1.5) (some really profound, others less), but most of which are simply called "Bernstein's inequality" or "Bernstein type inequality". A big treatise would not suffice to present all of them. I recall that about twenty years ago, around 1981, the late S. K. Pichorides (1940–1992) and myself thought of writing a fairly complete monograph *just on the sup-norm versions* of the Markov-Bernstein type inequalities, and in view of the enormity of the literature, we gave up.

This being said, what is the "most classical" statement (or, if we prefer, the "most widely known" statement of "Bernstein's inequality"? In other words, if we ask a random mathematical analyst to tell us quickly what is meant by "Bernstein's inequality", what is he/she likely to reply off the top of his/her head? In my opinion there are three such "likely" statements. The first is the above inequality (1.5) for all *real* trigonometric polynomials (1.4). The second is the extension of this same (1.5) to all *complex-valued* trigonometric polynomials of the form

$$f(t) = \sum_{k=-n}^{n} a_k e^{ikt} \quad (a_k \in \mathbb{C} \text{ for all } k = -n, \dots, n) \tag{1.9}$$

(with equality in (1.5) now reached if and only if $f(t) = ae^{int} + be^{-int}$ where a and b are complex constants.) The third is the following: If

$$P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X] \quad \text{and} \quad \|P\|_\infty := \max_{t\in\mathbb{R}} \left|P(e^{it})\right|,$$

then

$$\|P'\|_\infty \leqslant n\|P\|_\infty \tag{1.10}$$

(with equality in (1.10) if and only if $P(X) = aX^n$.)

Now, are the above three statements of "the classical Bernstein inequality" *equivalent*? (By "equivalent" I mean, in the present context, that each can be deduced from the others *in a quite easy manner*. Otherwise I am aware that any two true mathematical statements are equivalent!) The answer is that the first two are indeed equivalent, but that (1.10) is a weaker result. Indeed, the first statement is an obvious special case of the second, but the second is also a straightforward consequence of the first. (Here is a simple proof from G. G. Lorentz's book [23], by the same argument as the one presented right after (1.3) in this introduction, but I do not know who gave it first: With $f(t)$ of the form (1.9), select $\alpha \in \mathbb{R}$ so that $e^{i\alpha}f'(t)$ attains the value $\|f'\|_\infty$, say for $t = t_0$. Now $g(t) := \Re(e^{i\alpha}f(t))$ is of the form (1.4), so

$\|f'\|_\infty = e^{i\alpha} f'(t_0) = g'(t_0) \leqslant \|g\|_\infty \leqslant \|f\|_\infty$. The equality case is ignored in [23], but this is not hard either.) Now the third statement (1.10) is obviously a consequence of the second, and it is also (easily) derived *directly* from the first on page 45 of Bernstein's book [4]. Yet I do not know of any simple way of deriving the first or second statement from (1.10). On the other hand (1.10), as well as its extensions to all L^p norms with $p \geqslant 1$, have direct proofs much simpler than any of those available for the first two statements. The reader can work them out herself or look them up, *for example*, in [32]. Here is yet another way of seeing that the first two statements (1.5) are stronger than (1.10). Indeed they imply

$$|P'(e^{it})| \leqslant \frac{n}{2} \left(\|P\|_\infty + |P(e^{it})| \right) \tag{1.11}$$

(as easily checked, or see for example [17]), and (1.11) is obviously a deeper form of (1.10).

Does Markov's inequality (1.1) compare, depth-wise, to any ("classical" form of) Bernstein's inequality, although it was discovered utterly *independently* and twenty-two years earlier, as we saw? Using Bernstein's original theorem (1.3) and a nice theorem of Schur (Theorem 6, Chapter 3 of Lorentz's book [23]), we see that Markov's inequality is a corollary to (1.3) and thus to (1.5) for all *cosine* polynomials, but modulo Schur's theorem which is not trivial (although not very difficult either.) Altogether we may rate Markov's inequality (1.1) at about the same level of depth as Bernstein's inequality (1.5) for all trigonometric polynomials of the form (1.4) or (1.9).

The above inequalities of Markov and Bernstein (together with their equality cases) are typical examples of extremal problems, as they are equivalent to the problems of finding the maximum of $\|P'\|_{[-1,1]}/\|P\|_{[-1,1]}$ and of $\|f'\|_\infty/\|f\|_\infty$ where $\deg P \leqslant n$ and $\deg f \leqslant n$, ($P \not\equiv 0$ and $f \not\equiv 0$). I chose (the earliest versions of) the Markov-Bernstein inequalities as typical illustrations of extremal problems for two main reasons: 1) They were among the first (but certainly not *the very* first) that marked the beginning of the century-old "golden age" (twentieth century.) 2) They have been the starting point of an enormous literature concerning their generalizations, extensions, refinements and applications. Their usefulness in numerous areas of mathematics, physics and engineering has been extraordinary, ever since they were discovered. See *for example* [1], [8], [32] and many other good references.

This being said, there are at least two different meanings to the expression "extremal problem":

1. One meaning is obviously the *search* for (attained or unattained, global or local) suprema and/or infima of real-valued functions defined on a set (of functions, of polynomials, of numbers, of measures, etc.). With this first acceptation, there is a considerable overlap between the subject of "extremal problems" and that of "inequalities" (or rather "*optimal* inequalities").
2. Another one is a problem *pertaining to properties* of such suprema and infima (whether local or global, etc.), for example that of their distribution. This will not be discussed in this paper, but hopefully will be in [39].

The Bernstein-Markov inequalities are typical examples of polynomial extremal problems which find their roots in the nineteenth century theory of approximation and interpolation. But the majority of twentieth century polynomial extremal problems find their roots in the nineteenth century theory of trigonometric and Fourier series. We will (briefly) mention a few examples in this paper. The reader should, however, first read the next section on "peculiarities" of this paper.

1.2. *Peculiarities and aims of this paper*

I first explain in a page or two why this paper is a quite peculiar one. Its aims will then follow logically.

In May 2000, a few weeks before this July 2000 ASI, I realized that my original title "*Extremal problems on polynomials and trigonometric polynomials—a century of progress*" was too ambitious and unrealistic for a $2\frac{1}{2}$-hour tutorial lecture. So I toned it down to the present title. As for my oral lecture, I personally found it unsatisfactory, but in view of the audience's enthusiasm it seemed successful. I do not consider this to be my personal success but rather the merit of this delightful subject itself, where the problems and theorems are most often simple to state but the proofs are *sometimes* very hard and profound, sometimes easy and straightforward. I hope to return to this "instability phenomenon" of mathematical statements (with its sometimes very sad consequences) either at the end of this paper, or more likely in [39].

I intended to write up my lecture for the ASI proceedings but a major illness and accident in my family left me no free time in August and September of 2000. Thus I missed the submission deadline and decided to write instead a much longer version, perhaps a whole monograph. But J.-P. Kahane (who was my teacher and whose judgment I respect) pointed out that a monograph on such a huge subject might take years for a single person, and in addition the author might be tempted to stop the write-up and work on the fascinating open problems of this subject. So Kahane advised me to request a deadline extension from the editor and try to write the ASI paper anyway.

As the editor granted such extension, I could use a week or so of peace, in the house of a friend in Newport, Rhode Island, to produce the present paper at the beginning of November 2000. So this text was written within a very short period, when I had no access to a library nor to my personal notes and documents either (as I had left nearly all of them in Europe), and also the editor had imposed a 30-page bound on the length of the paper. Thus I had no choice but to write a "highly imperfect" paper, with the following features (some of which the reader might interpret as the "aims" of this paper):

1. A number of topics are *touched upon* but all of them very briefly, and most of them quite superficially. But I intend to write soon an extended version [39] of this paper.
2. All of the topics are *among* those I had wished to mention anyway, but the "choice" of them (in this paper) is nowhere near any optimal order of priority. The choice of topics here is by no means the result of thoughtful considerations, but rather the result

of time limitations only: To write this paper in the short alloted time, I had to write down whatever would come to my mind after some very quick thinking, in a nearly random way. Had I written this paper one week earlier or one week later, the choice of topics would have been quite different. However, in the extended version [39], not only the same topics will be treated in more detail and depth, but also other topics will be added.

3. Although many nice topics are not even mentioned here, all those touched upon in this paper are subjects that, I for one, do consider extremely interesting. All of them can be described as "Erdös-style harmonic analysis". Proofs are omitted (save for a couple exceptions), due to time and space limitations. Many of the topics mentioned concern my own results and conjectures, usually unpublished (and even unwritten.) This is again simply due to time limitations: After all, my own results (whether written or not) where the most accessible ones under such "space-time" pressure!
4. No systematic history of any of the topics mentioned will be presented in this paper, but hopefully this will be done (to some extent) in the extended version [39], at least for some of the topics. However, a number of historical points or anecdotes will be encountered here. One of my personal obsessions is *historical accuracy and honesty*, although I am not always successful at that. It is a very unfortunate fact that historical accuracy and honesty is *not* widespread among mathematicians, whether due to ignorance, negligence or malevolence. Authors *very frequently* reproduce, in their publications, historical errors or lies they have read elsewhere, without checking anything, thus perpetuating errors or lies. Such sad things are often unintentional, but too often intentional also. I have had my fair share of this kind of historical negligence, but it was always unintentional.
5. What was just said about historical points can also be said about *references*, and the way erroneous or inappropriate references are perpetuated and carried from publication to publication. Despite my good intentions in this respect, I cannot guarantee that all the references in *this* paper are correct, in view of the circumstances under which this paper was written. Quite often, I do not even give any reference at all to a result I am quoting only by relying on my memory, as I do not have the reference on my desk at the moment of writing. (This is the case even with *my own* papers!)
6. In conclusion, this paper is not to be viewed as a survey of extremal problems, even in a very limited sense. Its main function is to serve as a "memorandum" for myself with a view to the extended version [39]. The reader's criticisms, comments (and whatever information he/she could give me) are welcome and will be gratefully acknowledged.

2. Back to Bernstein's inequality

Two of the most useful refinements of Bernstein's inequality are Szegö's inequality [42] proved in 1928 and the Schaake-van der Corput inequality [40] proved in 1935. The latter was quickly noted to be a weaker form (and a corollary) of Szegö's inequality, but has a

usefulness of its own. Both are among the very few results that I will prove in this paper. These proofs will not be those of the original authors, as they were quite complicated, but proofs of my own (as they are the simplest ones I know of.)

2.1. *Statement of the Schaake-van der Corput inequality*

THEOREM 2.1 (Schaake & van der Corput [40]).
For any real trigonometric polynomial

$$f(t) = A_0 + \sum_{k=1}^{n} (A_k \cos kt + B_k \sin kt) \tag{2.1}$$

we have, for all $t \in \mathbb{R}$, the following refinement of Bernstein's inequality (1.5):

$$(f'(t))^2 + n^2 (f(t))^2 \leqslant n^2 \|f\|_\infty^2. \tag{2.2}$$

In other words, we have the identity

$$\|f'^2 + n^2 f^2\|_\infty = n^2 \|f\|_\infty^2. \tag{2.3}$$

Remarks on the equality case of (2.2). Although the inequality (2.2) is a refinement of Bernstein's inequality (1.5), the *equality cases* of these two inequalities are very different in nature. For Bernstein's inequality (1.5), the assumption that $|f'(t_0)| = n\|f\|_\infty$ for *at least one* point $t_0 \in \mathbb{R}$ implies that $f(t)$ is a sinusoidal function: $f(t) = A\cos nt + B\sin nt$. But for any (arbitrary) real trigonometric polynomial $f(t)$ of the form (2.1), the Schaake-van der Corput inequality (2.2) is obviously always an equality at every $t_0 \in \mathbb{R}$ where $|f(t_0)| = \|f\|_\infty$. Also (2.2) is an equality at *every* $t \in \mathbb{R}$ whenever $f(t) \equiv A\cos nt + B\sin nt$ or $f(t) \equiv$ constant, and the converse of this is also trivially true. However one can easily show that if (2.2) is an equality for at least $2n$ distinct points of the semi-open interval $[0, 2\pi[$, then either $f(t) \equiv A\cos nt + B\sin nt$ or $f(t) \equiv$ constant. In this last remark the number $2n$ is *optimal*, (*i.e.*, minimal): One can find non-sinusoidal and non-constant real trigonometric polynomials of the form (2.1) for which (2.2) is an equality at $2n-1$ distinct points of the semi-open interval $[0, 2\pi[$.

2.2. *Saffari's proof of the Schaake-van der Corput inequality*

We will obtain (2.2) as a corollary to Bernstein's inequality (1.5) *together with its equality case*, that is, $\|f'\|_\infty = n\|f\|_\infty$ *if and only if* $f(t)$ is a sinusoidal function of the form

$$f(t) = A\cos nt + B\sin nt. \tag{2.4}$$

Our proof is by contradiction. Let V_n denote the vector space (on the field $\mathbb{R}$) of all real trigonometric polynomials of the form (2.1), with degree $\leqslant 2n$. For any $f \in V_n \backslash \{0\}$ (*i.e.*,

$f \not\equiv 0$), define:

$$H(f) := \frac{\|f'^2 + n^2 f^2\|_\infty}{n^2 \|f\|_\infty^2} \tag{2.5}$$

and let

$$H_{\max} := \max_{f \in V_n \backslash \{0\}} H(f) \tag{2.6}$$

To prove that the maximum $H_{\max}$ defined by (2.6) indeed exists, first note that V_n is a finite-dimensional space over $\mathbb{R}$ (since $\dim_{\mathbb{R}} V_n = 2n+1$). If V_n is equipped with the norm topology, then (2.5) defines a *continuous* mapping H from $V_n \backslash \{0\}$ into $\mathbb{R}$: Indeed, since $\dim_{\mathbb{R}} V_n$ is finite, the (linear) differentiation operator $f \to f'$ is continuous on V_n, and so the mapping $f \to f'^2 + n^2 f^2$ is a continuous function from V_n into V_{2n-1}. Also all the norms (and in particular the sup-norm function $\varphi \to \|\varphi\|_\infty$) are continuous on the finite-dimensional vector spaces V_n and V_{2n-1}, ($\dim_{\mathbb{R}} V_{2n-1} = 4n-1$), so the mapping H defined by (2.5) is continuous as claimed. Also $H(\lambda f) = H(f)$ for all $\lambda \in \mathbb{R}$ with $\lambda \neq 0$, hence the maximum defined by (2.6) is the same as the maximum of $H(f)$ when f is restricted to the (compact) unit sphere of V_n for the sup-norm metric, and this last maximum is indeed attained.

Obviously $H_{\max} \geqslant 1$ (since $H(f) \geqslant 1$ for every $f \in V_n \backslash \{0\}$), and the Schaake-van der Corput inequality we wish to prove is equivalent to the equality $H_{\max} = 1$. To prove this by contradiction, we suppose henceforth that

$$H_{\max} > 1. \tag{2.7}$$

Let $\psi \in V_n \backslash \{0\}$ be some trigonometric polynomial of the form (2.1) for which

$$H(\psi) = H_{\max}. \tag{2.8}$$

We first note that ψ is not a constant function, since otherwise $\psi' \equiv 0$ and so, by (2.5), $H(\psi) = 1$ contrary to our assumption (2.7). Thus $\psi' \not\equiv 0$, *i.e.*, $\psi' \in V_n \backslash \{0\}$, hence we can consider

$$H(\psi') = \frac{\|\psi''^2 + n^2 \psi'^2\|_\infty}{n^2 \|\psi'\|_\infty^2}. \tag{2.9}$$

Let the trigonometric polynomial $(\psi'(t))^2 + n^2 (\psi(t))^2$ reach its absolute maximum at $t_0 \in \mathbb{R}$:

$$(\psi'(t_0))^2 + n^2 (\psi(t_0))^2 = \|\psi'^2 + n^2 \psi^2\|_\infty^2. \tag{2.10}$$

So its derivative vanishes at t_0:

$$2\psi'(t_0)\psi''(t_0) + 2n^2 \psi(t_0)\psi'(t_0) = 0. \tag{2.11}$$

By (2.9) we cannot have $\psi'(t_0) = 0$, otherwise $\|\psi'^2 + n^2\psi^2\|_\infty = n^2\left(\psi(t_0)\right)^2$ by (2.10), and so we would have

$$H(\psi) = \frac{\|\psi'^2 + n^2\psi^2\|_\infty}{n^2\|\psi\|_\infty^2} = \frac{n^2\left(\psi(t_0)\right)^2}{n^2\|\psi\|_\infty^2} \leqslant 1$$

which is impossible in view of (2.7) and (2.8). Thus $\psi'(t_0) \neq 0$, so upon dividing both sides of (2.11) by $\psi'(t_0)$ we get

$$\psi''(t_0) + n^2\psi(t_0) = 0. \tag{2.12}$$

Note, incidentally, that our observation $\psi'(t_0) \neq 0$ is another way of seeing that $\psi' \not\equiv 0$ so as to justify (2.9). Finally, since for sinusoidal functions $f(t)$ of the form (2.4) we obviously have $H(f) = 1$, it follows from (2.7) and (2.8) that $\psi(t)$ is *not* a sinusoidal function of the form (2.4). Therefore, by the equality case of Bernstein's inequality,

$$\|\psi'\|_\infty < n\|\psi\|_\infty. \tag{2.13}$$

We now have all we need to find a lower bound for (2.9) which will yield the desired contradiction to (2.7). By (2.9) and the *strict* inequality (2.13),

$$\begin{aligned}
H(\psi') &> \frac{\|\psi''^2 + n^2\psi'^2\|_\infty}{n^4\|\psi\|_\infty^2} \geqslant \frac{\left(\psi''(t_0)\right)^2 + n^2\left(\psi'(t_0)\right)^2}{n^4\|\psi\|_\infty^2} \\
&= \frac{n^4\left(\psi(t_0)\right)^2 + n^2\left(\psi'(t_0)\right)^2}{n^4\|\psi\|_\infty^2} \qquad \text{[by using (2.12)]} \\
&= \frac{n^2\left(\psi(t_0)\right)^2 + \left(\psi'(t_0)\right)^2}{n^2\|\psi\|_\infty^2} \qquad \text{[upon dividing by } n^2\text{]} \\
&= \frac{\|\psi'^2 + n^2\psi^2\|_\infty}{n^2\|\psi\|_\infty^2} \qquad \text{[by (2.10)]} \\
&= H_{\max}. \qquad \text{[by (2.8)]}
\end{aligned}$$

Thus $H(\psi') > H_{\max}$, which contradicts the maximality of $H_{\max}$ defined by (2.6). Therefore our assumption (2.7) leads to a contradiction, and thus we have $H_{\max} = 1$ and have proved the Schaake-van der Corput inequality. □

2.3. *Other proofs or reformulations of the Schaake-van der Corput inequality*

I gave my own proof of the Schaake-van der Corput inequality as it is one of the simplest ones I know, and certainly simpler than the original one [40]. Just relying on my memory at the moment of this write-up, I think I know a good dozen different proofs of this inequality. There might be many more in the literature. I hope to give some of them in [39]. In this connection, here is an unfortunate anecdote: Just over twenty years ago, in the late 1970's, I found my above proof of the Schaake-van der Corput inequality and gave a seminar talk

about it at Orsay. The following week the late S. K. Pichorides (1940-1992) talked in the same seminar on a nice application of the Schaake-van der Corput inequality. Shortly after, without Pichorides or myself knowing about it, two *physicists* happened to submit a joint mathematical note to the French "Comptes Rendus de l' Académie des Sciences" giving a beautiful number-theoretic proof of the Schaake-van der Corput inequality (which they believed to be a new result), and they had applications of this inequality to *optics*. Their note was abruptly rejected by the referees (two mathematicians at Orsay, none of whom were J.-P. Kahane or Y. Meyer), with the comment: "*This result is not new, it is a recent theorem of Saffari and Pichorides*". So much for the competence of such "referees", who never bothered to ask me or Pichorides. Only one year after, when I heard about this rejection, could I tell those "referees" that the result actually went back to 1928 [42] but such a nice new proof should not have been rejected. In [39] I will present the (unpublished) proof of those physicists, and disclose their names.

Polynomial reformulation of the Schaake-van der Corput inequality (2.2). Let the polynomial $P(x) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$ be self-inversive, that is, $a_{n-k} = \overline{a}_k$ for all $k = 0, 1, \ldots, n$. Then the Bernstein inequality $\|P'\|_\infty \leqslant n\|P\|_\infty$ can be improved to:

$$\|P'\|_\infty = \frac{n}{2}\|P\|_\infty \qquad \textit{(yes, equality!)} \tag{2.14}$$

Proof. The truth of (2.14) when n is *odd* follows from the truth of (2.14) when n is *even* by considering the even-degree polynomial $P(z^2)$. So, without loss, we may suppose n even, $n = 2m$. The (real-valued) trigonometric polynomial $f(t) := e^{-imt}P(e^{it})$ is then of the form (2.1), with degree m, hence upon writing $P(e^{it}) = e^{imt}f(t)$ and differentiating both sides, we get

$$ie^{it}P'(e^{it}) = e^{imt}\left(f'(t) + imf(t)\right),$$

hence

$$\left|P'(e^{it})\right|^2 = \left(f(t)\right)^2 + m^2\left(f(t)\right)^2 . \tag{2.15}$$

Thus (2.14) is indeed equivalent to (2.3). □

The above polynomial reformulation (2.14), like the original formulation (2.2), was re-discovered by many people. Some of these proofs do shed new light on the subject, see [39].

2.4. *"Self-improvement" of the Schaake-van der Corput inequality*

This is an important topic, and we will see an application of it in Section 4. Let the real trigonometric polynomial $f(t)$ be as in (2.1), and consider the trigonometric polynomial (of degree $\leqslant 2n-1$):

$$f_1(t) := \left(f'(t)\right)^2 + n^2\left(f(t)\right)^2 . \tag{2.16}$$

By (2.3) we have $\|f_1\|_\infty = n^2\|f\|_\infty^2$. Applying the Schaake-van der Corput inequality (2.2) to this $f_1(t)$ (with n replaced here by $2n-1$), we obtain

$$(f_1'(t))^2 + (2n-1)^2 (f_1(t))^2 \leqslant (2n-1)^2\|f_1\|_\infty^2 \tag{2.17}$$

which, on dividing by $(2n-1)^2$ and using (2.16), can be rewritten:

$$\frac{4\,(f'(t))^2}{(2n-1)^2}\left(f''(t)+n^2 f(t)\right)^2 + \left((f'(t))^2 + n^2\,(f(t))^2\right)^2 \leqslant n^4\|f\|_\infty^4. \tag{2.18}$$

This is a refinement of (2.2). We can continue such improvements indefinitely, but they become too complicated and, mostly, this method is very *wasteful* as it ignores the fact that $f_1(t) \geqslant 0$ while for *non-negative* trigonometric polynomials there is a better form of the Schaake-van der Corput inequality than (2.2), namely inequality (2.19) below:

THEOREM 2.2 (The case of non-negative trigonometric polynomials).
If a trigonometric polynomial $g(t)$ of the form (2.1) satisfies $g(t) \geqslant 0$ for all $t \in \mathbb{R}$, then for all $t \in \mathbb{R}$

$$(g'(t))^2 + n^2\,(g(t))^2 \leqslant n^2\|g\|_\infty \cdot g(t). \tag{2.19}$$

Proof. Apply (2.2) to the polynomial $f(t) := g(t) - \frac{1}{2}\|g\|_\infty$. □

Now, applying (2.19) to the non-negative trigonometric polynomial $f_1(t)$ given by (2.16), we obtain for *any* (not necessarily non-negative) real trigonometric polynomial of the form (2.1):

$$\frac{4\,(f'(t))^2}{(2n-1)^2}\cdot\frac{(f''(t)+n^2 f(t))^2}{(f'(t))^2+n^2\,(f(t))^2} + (f'(t))^2 + n^2\,(f(t))^2 \leqslant n^2\|f\|_\infty^2 \tag{2.20}$$

which is a finer improvement of the Schaake-van der Corput inequality than (2.18). In turn, there is yet an improvement of (2.20) for non-negative trigonometric polynomials. Actually there is a rather nice method for refining (2.20) *indefinitely* while obtaining expressions which are not too complicated and also, in a sense, optimal. We will drop this matter here but might return to it in [39] under the (weird-sounding but appropriate) term of "*analytic bootstrapping*". Let us also note that, just as for (2.14), the other variations and refinements of the Schaake-van der Corput inequality can be reformulated in terms of algebraic polynomials, *whether self-inversive or not*. Thus, for $P(X) = \sum_{k=0}^n a_k X^k \in \mathbb{C}[X]$ *not necessarily self-inversive*,

$$\frac{4}{n^2}\left(\frac{d}{dt}\left|P(e^{it})\right|\right)^2 + \left|P(e^{it})\right|^2 \leqslant \|P\|_\infty^2. \tag{2.21}$$

Proof. Divide both sides of (2.19) by $n^2 g(t)$ and take $g(t) = |P(e^{it})|^2$. □

2.5. *Elementary applications of the Schaake-van der Corput inequality*

If $f(t)$ is a real trigonometric polynomial of the form (2.1), then one can refine the trivial inequality $\|f\|_2 \leqslant \|f\|_\infty$ into:

$$n^{-2}\|f'\|_2^2 + \|f\|_2^2 \leqslant \|f\|_\infty^2. \tag{2.22}$$

This is obtained by integrating (2.2) on $[0, 2\pi]$. □

Although we lost information in the integration process, and (2.22) is therefore weak and wasteful, it is still useful and I hope to give some application of (2.22) in [39]. An even more useful inequality is the following:

THEOREM 2.3. *If $g(t)$ is a non-negative trigonometric polynomial of the form (2.1), then one can refine the trivial inequality $\|g\|_2^2 \leqslant \|g\|_1 \cdot \|g\|_\infty$ into:*

$$n^{-2}\|g'\|_2^2 + \|g\|_2^2 \leqslant \|g\|_1 \cdot \|g\|_\infty. \tag{2.23}$$

Equivalently, if $P(X) = \sum_{k=0}^n a_k X^k \in \mathbb{C}[X]$ is a (not necessarily self-inversive) polynomial with complex coefficients, then the trivial inequality $\|P\|_4^2 \leqslant \|P\|_2^2 \cdot \|P\|_\infty^2$ can be refined into:

$$n^{-2}\left\|\frac{d}{dt}\left(|P(e^{it})|^2\right)\right\|_2^2 + \|P\|_4^4 \leqslant \|P\|_2^2 \cdot \|P\|_\infty^2. \tag{2.24}$$

Proof. To obtain (2.23) integrate (2.19) on $[0, 2\pi]$. Then take $g(t) = |P(e^{it})|^2$ to obtain (2.24). □

2.6. *Szegö's inequality*

One year after the appearance of their 1935 paper [40] in which (2.2) was proved, Schaake and van der Corput pointed out in a short note [41] that their result (2.2) was already contained in (*i.e.*, was a weaker form of) a more precise inequality published by Szegö [42] in 1928. The statement of Szegö's inequality is as follows:

THEOREM 2.4. *Let $f(t)$ be a real trigonometric polynomial of the form (2.1), and let $\tilde{f}(t)$ denote its conjugate trigonometric polynomial. Then, for all $t \in \mathbb{R}$,*

$$|nf(t) - \tilde{f}'(t)| + \sqrt{(f'(t))^2 + \left(\tilde{f}'(t)\right)^2} \leqslant n\|f\|_\infty\,. \tag{2.25}$$

In this paper I will not give Szegö's original proof of (2.25) but, in section 2.7 below, I will give the (possibly new) simple proof *by myself and the late S. K. Pichorides* [36], via an interpolation method. Yet it is conceivable that the Saffari-Pichorides proof of (2.25) (or a similar one) is already in the literature.

A few remarks before quitting this section 2.6. First, in (2.25) $\tilde{f}'(t)$ is the conjugate function of the derivative $f'(t)$ and also the derivative of the conjugate function $\tilde{f}(t)$, as these are equal. So, no ambiguity here.

Another remark is that Szegö's inequality (2.25) is finer than the Schaake-van der Corput inequality (2.2) simply because, for all $t \in \mathbb{R}$,

$$\sqrt{(f'(t))^2 + n^2 (f(t))^2} \leqslant \left| nf(t) - \tilde{f}'(t) \right| + \sqrt{(f'(t))^2 + \left(\tilde{f}'(t)\right)^2}. \tag{2.26}$$

To get (2.26), take $x = nf(t)$, $y = f'(t)$, $z = \tilde{f}'(t)$ in the inequality

$$\sqrt{x^2 + y^2} \leqslant |x - z| + \sqrt{y^2 + z^2} \tag{2.27}$$

which holds for all $x, y, z \in \mathbb{R}$ and is just the ordinary triangle inequality for Euclidean norm.

Equality cases of Szegö's inequality (2.25): One can see in several ways (***e.g.***, from our proof in the next section 2.7) that equality holds *for all* $t \in \mathbb{R}$ if and only if $f(t)$ is of the form

$$f(t) = A_0 + A \cos nt + B \sin nt, \tag{2.28}$$

i.e., for a wider class than the ones for the Schaake-van der Corput and Bernstein inequalities. Also, for a given $f(t)$, there are points $t_0 \in \mathbb{R}$ (depending on f) where equality holds in (2.25). This discussion is interesting but we drop it in this paper.

A final remark: By throwing out $\left| nf(t) - \tilde{f}'(t) \right|$ in (2.25), one obtains:

$$(f'(t))^2 + \left(\tilde{f}'(t)\right)^2 \leqslant n^2 \|f\|_\infty^2 \tag{2.29}$$

or, *equivalently*,

$$\|P'\|_\infty \leqslant n \|\Re(P)\|_\infty \tag{2.30}$$

for all polynomials $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$. Both (2.29) and (2.30) had, of course, been noted by Szegö (in [42] and elsewhere). They are also in Zygmund's classical treatise [44], together with their L^p generalizations, ($p \geqslant 1$), and therefore well known to a wide public of present-day analysts. However, (2.2) (which, incidentally, does not imply (2.29) nor is implied by it either) and (2.25) are *not* in Zygmund's treatise [44] and thus are not as widely known as they deserve. In [44], Zygmund does attribute (2.29) to Szegö and refers to Szegö's 1928 paper [42], yet the proof of (2.29) presented by Zygmund in [44] is *not* Szegö's original proof but another proof akin to one of the earliest proofs of Bernstein's inequality (1.5) by Marcel Riesz [35] (or to the interpolation method of section 2.7 below). The last time I saw the late Professor Zygmund, in 1984, I asked him why he chose to include (2.29) in his treatise [44] *but not* the similar (and even more useful) Schaake-van der Corput inequality (2.2) nor the strong Szegö inequality (2.25) either. Zygmund replied that, until my conversation with him about this matter, he had never heard of the inequalities (2.2) and (2.25), and although he had quoted Szegö's paper [42] in his treatise [44], he actually had never taken a look at Szegö's paper [42] (where (2.25) was originally proved). He had

also not read some other papers quoted in his treatise [44], and in which (2.2) and (2.25) were mentioned together with some extensions. At least Zygmund was perfectly honest and candid in this respect. He was a good friend of Szegö, and the two did great joint research (some of which was on polynomials, see [43]).

2.7. *The Pichorides-Saffari proof of Szegö's inequality*

(Again, I note that our proof below might already be in the literature.)

The proof is based on the identity

$$\sum_{h=1}^{n} F_{n-1}\left(\frac{2h\pi-\alpha}{n}\right) e^{ik(2h\pi-\alpha)/n} = n-k+ke^{-i\alpha} \tag{2.31}$$

for all $\alpha \in \mathbb{R}$ and all integers k, n such that $0 \leqslant k \leqslant n$, where $F_{n-1}(x)$ denotes the Fejér kernel of degree $n-1$:

$$F_{n-1}(x) := \sum_{r=-n}^{n} \left(1-\frac{|r|}{n}\right) e^{irx} = \frac{1}{n}\cdot\left(\frac{\sin\frac{nx}{2}}{\sin\frac{x}{2}}\right)^2 . \tag{2.32}$$

To check (2.31), use the middle sum in (2.32) with $x=(2h\pi-\alpha)/n$, plug this expression into the left hand side of (2.31) and change the order of summation, then note that the new inner sum is always zero except for $r=-k$ and for $r=n-k$.

Another form of (2.31) is

$$\sum_{h=1}^{n} F_{n-1}\left(\frac{2h\pi-\alpha}{n}\right) e^{ik(2h\pi-\alpha)/n} = n-|k|+|k|\cos\alpha - ik\sin\alpha \tag{2.33}$$

for all $\alpha \in \mathbb{R}$ and all integers k, n such that $-n \leqslant k \leqslant n$.

[For $k \geqslant 0$, (2.33) is the same as (2.31); for $k<0$ take the complex conjugates on both sides of (2.31)].

On taking $k=0$, we have the following useful identity:

$$\sum_{h=1}^{n} F_{n-1}\left(\frac{2h\pi-\alpha}{n}\right) = n \qquad \text{(for all } \alpha\in\mathbb{R}\text{)}. \tag{2.34}$$

Now, for any integer k with $1 \leqslant k \leqslant n$, multiply both sides of (2.31) by $e^{i(kt+\varphi_k)}$ (with $t, \varphi_k \in \mathbb{R}$) and then take the real parts, to get

$$\sum_{h=1}^{n} F_{n-1}\left(\frac{2h\pi-\alpha}{n}\right)\cos\left(kt+\varphi_k+k\cdot\frac{2h\pi-\alpha}{n}\right) =$$

$$= (n-k)\cos(kt+\varphi_k)+k\cos(kt+\varphi_k-\alpha). \tag{2.35}$$

We now have all we need to prove the *interpolation formula*:

$$nf(t) - \tilde{f}'(t) + \tilde{f}'(t)\cos\alpha - f'(t)\sin\alpha = \\ = \sum_{h=1}^{n} F_{n-1}\left(\frac{2h\pi-\alpha}{n}\right) f\left(t+\frac{2h\pi-\alpha}{n}\right) \tag{2.36}$$

for any real trigonometric polynomial $f(t)$ of the form (2.1). Indeed, write $f(t)$ in the more convenient form

$$f(t) = A_0 + \sum_{k=1}^{n} r_k \cos(kt+\varphi_k), \qquad (r_k = \sqrt{A_k^2+B_k^2}) \tag{2.37}$$

and note that, by linearity of the differentiation $f \to f'$ and of the Hilbert transform $f \to \tilde{f}$, it suffices to check (2.36) for $f(t) \equiv 1$ and for $f(t) \equiv \cos(kt+\varphi_k)$, $(1 \leqslant k \leqslant n)$. For $f(t) \equiv 1$, (2.36) reduces to (2.34). For $f(t) \equiv \cos(kt+\varphi_k)$, we have $f'(t) = -k\sin(kt+\varphi_k)$ and $\tilde{f}'(t) = k\cos(kt+\varphi_k)$, so that (2.36) reduces to (2.35). Thus (2.36) is proved.

The desired inequality (2.25) follows from (2.36). Indeed by (2.32) (which implies $F_{n-1}(x) \geqslant 0$) and (2.34), for any fixed α and $t \in \mathbb{R}$, the modulus of the right-hand side of (2.36) is majorized by

$$\sum_{h=1}^{n} F_{n-1}\left(\frac{2h\pi-\alpha}{n}\right) \cdot \|f\|_\infty = n\|f\|_\infty .$$

Now, for any fixed $t \in \mathbb{R}$, the maximum modulus of the left-hand side of (2.36) (as α varies) is

$$|nf(t) - \tilde{f}'(t)| + \sqrt{\left(\tilde{f}'(t)\right)^2 + \left(f'(t)\right)^2}$$

and the proof of (2.25) is complete. □

2.8. *Some sources*

The very important subject of Markov-Bernstein type inequalities deserves a reasonably complete monograph. Pending the writing of such a book, which is no easy task, here are just a few good sources of information (*among many others*):

1. The 1983 book "Les inégalités de Markoff et de Bernstein" by Q. I. Rahman and G. Schmeisser [31] and its good *list of references*. This (rather short) mimeographed book is written in a delightfully reader-friendly style reminiscent of those of S. N. Bernstein [4] and C. de La Vallée Poussin [12]. A modest knowledge of French should be enough for reading this book.
2. P. Borwein and T. Erdelyi (either separately, or together, or with other co-authors) have many publications on this subject (and, by the way, on several other topics pertaining to the themes of this paper and to those of [39]). Most preprints of their publications can

be downloaded from the authors' well organized home pages at the Web sites of Simon Fraser University [9] and Texas A&M University [14], respectively. Although these reprints are not always the final forms of the publications and do sometimes contain typos, the papers are of a high level and also good sources of references (modulo a few errors).

3. Many Ph. D theses and other publications on this subject produced in the province of Québec, Canada, where there is a tradition of this type of studies. I have lost track of what is being done there, but the reader could look at the university Web sites and at the publication lists of "Presses de l'Université de Montréal.". One example (among others) is the 1983 Ph. D. thesis of C. Frappier [17].

3. Some types of flat polynomials

If a polynomial $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$ has at least two non-zero coefficients, then it cannot be "perfectly flat" (*i.e.*, have constant modulus) *on the whole unit circle*: The relation

$$|P(e^{it})| = \text{constant} \qquad (\text{for all } t \in \mathbb{R}) \tag{3.1}$$

is impossible. To see this, call $a_r X^r$ (resp. $a_s X^s$) the non-zero term of $P(X)$ of lowest (resp. highest) degree, ($0 \leqslant r < s \leqslant n$), and note that $|P(e^{it})|^2$ is a non-zero trigonometric polynomial with leading term $2|a_r a_s| \cos(mt + \beta - \alpha)$ where $m = s - r > 0$, $a_r = |a_r| e^{i\alpha}$, $a_s = |a_s| e^{i\beta}$.

However, for a variety of classes of non-monomial polynomials $P(X)$, the ideal (and impossible!) "perfect flatness" situation (3.1) can still be approximated in various ways. The diverse notions of "flatness" of a polynomial (or, more frequently, of *a sequence or a class of polynomials*) often refer to the various ways of approximating the ideal situation (3.1). For example, given a subset Γ (usually a subgroup) of the unit circle, if

$$|P(g)| = \text{constant} \qquad (\text{for all } g \in \Gamma) \tag{3.2}$$

then we say that the polynomial $P(X)$ is *perfectly flat* on the set Γ. There are some non-trivial open problems on that notion of perfect flatness, and we will briefly see a glimpse of them in section 3.1 below. Another example of a flatness requirement is to look for polynomials $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$ for which the moduli $|a_0|, |a_1|, \ldots, |a_n|$ are given non-negative numbers and for which the sup-norm $\|P\|_\infty := \max_{t \in \mathbb{R}} |P(e^{it})|$ is either as small as possible (*extremal problem*) or satisfies some smallness requirement. Thus, if we impose $|a_k| = 1$ for all $k = 0, 1, \ldots, n$ (or even without this restriction), the previous problem is equivalent to some smallness requirement on the "*crest factor*" (or "*peak factor*") $\|P\|_\infty / \|P\|_2$ where $\|P\|_2$ denotes the L^2-norm:

$$\|P\|_2 := \left(\frac{1}{2\pi} \int_0^{2\pi} |P(e^{it})|^2 dt \right)^{1/2} = \left(\sum_{k=0}^{n} |a_k|^2 \right)^{1/2} . \tag{3.3}$$

Actually, many flatness problems for polynomials can be expressed in terms of *comparison of two norms* in the vector space of polynomials of degree $\leqslant n$. We shall take brief looks at some examples in the sequel.

3.1. *Perfect flatness*

As we said above, a polynomial $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$ is "***perfectly flat***" on a subset Γ of the unit circle if (3.2) holds. It is easy to see that if Γ is a (finite or infinite) *subgroup* of the unit circle with $\operatorname{card}\Gamma > \deg P$ (where $\operatorname{card}\Gamma$ denotes the number of elements of Γ), then the constant in (3.2) equals $\|P\|_2$ defined by (3.3), but this is not always true for a subgroup Γ such that $\operatorname{card}\Gamma \leqslant \deg P$.

So the perfect flatness (3.2) cannot hold if Γ is the whole unit circle. Assuming $P(0) \neq 0$ and $\deg P = n$, it is easy to see that (3.2) cannot hold either if $\operatorname{card}\Gamma > 2n$, but that whenever $\operatorname{card}\Gamma \leqslant 2n$ there are some such $P(X)$ for which (3.2) holds.

An interesting *open* problem is the following: Given an integer $n \geqslant 1$, find the largest value $\Phi(n)$ of those integers $d \geqslant 1$ such that there exists a polynomial $P(X) = \sum_{k=0}^{n} a_k X^k \in \mathbb{C}[X]$ which is "***unimodular***" (***i.e.***, $|a_k| = 1$ for all $k = 0, 1, \ldots, n$) and for which the perfect flatness (3.2) holds on the group Γ_d of d-th roots of unity in $\mathbb{C}$. The computation of $\Phi(n)$ is easy for small n, for example: $\Phi(1) = 2$, $\Phi(2) = 4$, $\Phi(3) = 5$, $\Phi(4) = \Phi(5) = 6$. A list of values of $\Phi(n)$ has been computed (Björck & Saffari, 1998, and on-going search in 2000–2001). No values of n with $\Phi(n) \geqslant n + 3$ have been found so far, but we do not have enough evidence to conjecture that one always has either $\Phi(n) = n + 1$ or $\Phi(n) = n + 2$. The following partial results are known:

$$\Phi(n) \geqslant n + 1 \qquad \text{(for all } n \geqslant 1\text{)} \tag{3.4}$$

$$\Phi(n) \geqslant n + 2 \qquad \text{(for infinitely many } n\text{)} \tag{3.5}$$

$$\Phi(n) \leqslant 2n - 1 \qquad \text{(for all } n \geqslant 1\text{)}\,. \tag{3.6}$$

(3.4) follows from elementary results (Gauss sequences), see section 3.2 on bi-unimodular sequences. (3.5) is a non-trivial result of Björck and Saffari (unpublished). (3.6) is not hard to prove for *even* n, however for *odd* n it can be shown (Saffari, unpublished) to be equivalent to a profound theorem of Dresel, White and Hunt (see also the section in [39] on "Huffman sequences").

If, instead of polynomials with *complex* unimodular coefficients, we consider polynomials with *real* unimodular coefficients (***i.e.***, $a_k = \pm 1$ for all $k = 0, 1, \ldots, n$), then (3.2) becomes a different type of problem. It is very easy to check that, in that case, (3.2) never holds if $\operatorname{card}\Gamma \geqslant n + 3$. It can be proved (Saffari [38], unpublished) that (3.2) never holds either if $\operatorname{card}\Gamma = n + 2$. As for the impossibility of (3.2) for $\operatorname{card}\Gamma = n + 1$ (except for $n = 3$), it is *equivalent* to the famous "Hadamard Circulant Conjecture" due to Ryser [33], an open problem going back to the 1950's and stating that a circulant matrix of order L (with ± 1 entries) cannot be a Hadamard matrix *unless* $L = 4$.

3.2. *Bi-unimodular sequences*

This is a notion which we will define a little later and which will turn out to be *equivalent* to that of unimodular polynomials of degree n which are perfectly flat on the group Γ_{n+1} of $(n+1)$th roots of unity in $\mathbb{C}$ (see section 3.1 above). I mention this notion at this point because, historically, it can be considered as the oldest idea of "flat polynomials" since it goes back to Gauss. Indeed, let L be any *odd* integer $\geqslant 3$, and consider the "*Gaussian sequence*" of length L:

$$a_k := \omega^{\lambda k^2}, \quad (k = 0, 1, \ldots, L-1) \tag{3.7}$$

where $\omega = \exp(2i\pi/L)$ is the first primitive L-th root of 1 in $\mathbb{C}$, and λ is any integer relatively prime to L. Equivalently, we may consider the infinite sequence (a_k), $(k \in \mathbb{Z})$, as periodic of period L. Defining the normalized discrete Fourier transform (DFT) of a_k as the sequence

$$\hat{a}_r := \frac{1}{\sqrt{L}} \sum_{k=0}^{L-1} a_k \cdot \omega^{rk}, \quad (r = 0, 1, \ldots, L-1) \tag{3.8}$$

we can easily check that $(\hat{a}_r)$ is unimodular as well: $|\hat{a}_r| = 1$ for all $r = 0, 1, \ldots, L-1$. This fact was known to Gauss (to whom, by the way, the earliest ideas of Fourier Analysis can be tracked down, and not to Fourier or Clairaut). More generally, given any integer $L \geqslant 2$, whether odd or even, any finite sequence (a_k), $(k = 0, 1, \ldots, L-1)$, of L complex numbers will be called "*bi-unimodular*" if it has modulus one ($|a_k| = 1$ for all $k = 0, 1, \ldots, L-1$) and if its normalized DFT

$$\hat{a}_r := \frac{1}{\sqrt{L}} \sum_{k=0}^{L-1} a_k \cdot \omega^{rk}, \quad (\omega = e^{2i\pi/L};\ r = 0, 1, \ldots, L-1)$$

has modulus one, too: $|\hat{a}_r| = 1$ for all $r = 0, 1, \ldots, L-1$. (The term "bi-unimodular" was coined by G. Björck and myself in our 1995 joint paper [7]). Thus, for *odd* L, the Gaussian sequence (3.7) is an example of a bi-unimodular sequence of length L. The sequence (3.7) is *not* bi-unimodular when L is *even*, but in this case we have another type of Gaussian sequence which *is* bi-unimodular:

$$b_k := \xi^{\lambda k^2}, \quad (k = 0, 1, \ldots, L-1) \tag{3.9}$$

where $\xi = \exp(i\pi/L)$ is the first primitive root of unity of order $2L$ in $\mathbb{C}$, and λ again any integer relatively prime to L. Note that, while (3.7) is bi-unimodular for odd L but not for even L, similarly (3.9) is bi-unimodular for even L but not for odd L.

Obviously a unimodular sequence (a_k), $(k = 0, 1, \ldots, L-1)$, of length L is bi-unimodular if and only if its "associated polynomial" $P(X) = \sum_{k=0}^{n} a_k X^k$ (of degree $n = L-1$) is perfectly flat on the group Γ_L of L-th roots of 1 in $\mathbb{C}$.

The starting point of the theory of bi-unimodular sequences was an oral question asked by Per Enflo in 1983 at Stockholm University [13]: *If p is a given odd prime number, is it true*

that the Gaussian sequences $a_k = \omega^{\lambda k^2+\mu k}$, *(*$\omega = e^{2i\pi/p}$, λ *and* μ *integers with* p *not dividing* λ*), (*$k = 0, 1, \ldots, p-1$*), are the only unimodular sequences of length* p*, with* $a_0 = 1$*, whose normalized* DFT *has modulus one?*

In our present vocabulary, Per Enflo was asking whether such Gaussian sequences were the only normalized ($a_0 = 1$) bi-unimodular sequences of odd prime length. If the answer had been "yes", it would have helped him with estimations of some exponential sums. Later on he found another method (not requiring an answer to the above question) to carry out the estimations of his exponential sums anyway.

For $p = 3$ the answer is trivially "yes", and for $p = 5$ Lovász [24] checked that the answer is "yes" as well.

In 1984 G. Björck (Stockholm University) was trying, by computer search, to check that for $p = 7$ the answer to Per Enflo's question was "yes" as well, when suddenly the counter-example

$$(1, 1, 1, e^{i\theta}, 1, e^{i\theta}, e^{i\theta}) \qquad (\text{with } \theta = \arccos(-3/4)) \tag{3.10}$$

"popped out" (as Björck put it!). Later on in 1984 Björck found, again by computer search, other counter-examples to Per Enflo's question, including

$$(1, 1, e^{i\sigma}, 1, 1, 1, e^{i\sigma}, e^{i\sigma}, e^{i\sigma}, 1, e^{i\sigma}) \tag{3.11}$$

(with $\sigma = \arccos(-5/6)$). When early in 1985 Björck presented these counter-examples at the A. Haar memorial Conference [5], he still had no idea of the structure of the sequences (3.10) and (3.11). Actually, if in the sequence (3.10) [resp. (3.11)] we replace the first term by zero and the terms $e^{i\theta}$ [resp. $e^{i\sigma}$] by -1, we get the "*Legendre symbol*" sequence modulo 7 [resp. mod 11]: The terms $e^{i\theta}$ [resp. $e^{i\sigma}$] are located at the quadratic non-residues (modulo 7, resp. modulo 11). This is a fact that Björck observed a little later, in the fall of 1985, and that he subsequently generalized ([6], 1990) to every prime $\equiv -1 \pmod 4$:

If in the p-term *"Legendre symbol" sequence* $(0, 1, \ldots, -1)$*, (*p *any prime* $\equiv -1$ *modulo 4) we replace the first term zero by 1 and every* -1 *by*

$$\exp\left(i \arccos \frac{1-p}{1+p}\right) = \frac{1-p}{1+p} + i\frac{2\sqrt{p}}{1+p} \tag{3.12}$$

we obtain a bi-unimodular sequence of length p*, with only two values, namely 1 and the number given by (3.12).*

In sections 3.5 and 3.6 below I shall determine *all* bi-unimodular sequences with only two values: it will turn out, in particular, that Björck's above theorem is not specific to prime numbers but can be extended to all integers v (necessarily $\equiv -1$ modulo 4) for which there exists a so-called "Hadamard-Paley cyclic difference set." Thus it will be seen to work for $v = 15$, which is the smallest non-prime v with this property. From the same discussion of sections 3.5 and 3.6 it will follow that such a bi-unimodular sequence (with *only two values*) cannot exist for any length $\equiv 1 \pmod 4$. However, for *prime* lengths $\equiv 1 \pmod 4$, Björck proved (in the same paper [6]) the next best thing:

If in the p-term "Legendre symbol" sequence $(0, 1, \dots, -1, \dots, 1)$, ($p$ any prime $\equiv 1$ modulo 4) we replace the first term 0 by 1, every term 1 by

$$\eta := \exp\left(i \arccos \frac{\delta\sqrt{p}-1}{p-1}\right) = \frac{\delta\sqrt{p}-1}{p-1} + i\frac{\sqrt{p^2-3p+2\delta\sqrt{p}}}{p-1} \tag{3.13}$$

(with any choice of $\delta = \pm 1$) and every -1 by the complex conjugate $\overline{\eta}$, with the same choice of $\delta = \pm 1$, then we obtain a bi-unimodular sequence of length p (which thus has first term 1 and only two other values, namely η and $\overline{\eta}$).

In the 1990's more research was done on bi-unimodular sequences (Saffari [37], Björck & Saffari [7], Haagerup [19], ...) and there are some very nice results and open problems on this subject, that I hope to discuss in [39].

3.3. *Digression on (v, k, λ) difference sets*

This brief digression into Combinatorics (in this section and in next section 3.4) is just intended to recall, for the benefit of the harmonic analyst reader, some definitions and elementary facts which will be useful in the discussion (in sections 3.5 and 3.6) of polynomials whose coefficients only take *two* values (whether unimodular or not) and which are perfectly flat on the group Γ_L of L-th roots of 1 in $\mathbb{C}$, ($L = 1 + \deg P$).

There are currently two *entirely different and unrelated* types of mathematical objects, both carrying (unfortunately) the same name of "difference sets". We will be concerned only with the second notion, yet I will define both just to avoid any possible confusion (as I have often witnessed instances of such confusion between the two notions).

The set $\Delta(S) := \{x - y : x, y \in S\}$, where S is any subset of an (additive) abelian group G, is often called the "difference set" of S. This notion is well known to all analysts because of the result saying that if $G = \mathbb{R}^n$ and if S has positive Lebesgue measure, then the origin belongs to the interior of $\Delta(S)$ and therefore $\Delta(S)$ has non-empty interior. There is an enormous literature on this notion and its extensions, in such areas as analysis, algebra, combinatorics and number theory. I will not say anything else on this, as it is of no concern to us here.

The second notion of "difference sets" (the one which *does* interest us here) pertains to *finite* groups only. Let G be any finite group, abelian or not, with neutral element denoted by e. A subset D of G is called a left (resp. right) (v, k, λ) *difference set* if $\operatorname{card} G = v$, $\operatorname{card} D = k$ and the intersection $(uD) \cap D$ (resp. $(Du) \cap D$) has cardinality λ whenever $u \in G$, $u \neq e$. The term "difference set" is due to the fact that such sets were first considered in additive groups $\mathbb{Z}/n\mathbb{Z}$. Note that $D \subset G$ is a left (v, k, λ) difference set if and only if $D^{-1} := \{x^{-1} : x \in D\}$ is a right (v, k, λ) difference set. Also if D is a left (resp. right) (v, k, λ) difference set, then its complement $D' := G \backslash D$ is also a left (resp. right) (v', k', λ') difference set, with

$$v' = v, \quad k' = v - k, \quad \lambda' = v - 2k + \lambda. \tag{3.14}$$

The most obvious examples of difference sets are the following four types (called the "*trivial difference sets*"):

1. The empty set $\emptyset$, which is a $(v, 0, 0)$ difference set.
2. Singletons, which are $(v, 1, 0)$ difference sets.
3. $D = G$, which is a (v, v, v) difference set.
4. Complements of singletons, which are $(v, v-1, v-2)$ difference sets (if $v \geqslant 2$).

The simplest (and, historically, the first) example of a *non-trivial* difference set is due to Paley [29]: *If p is a prime $\equiv -1$ (mod4), then the set of all (non-zero) quadratic residues modulo p, and also the set of all quadratic non-residues modulo p, are (v, k, λ) difference sets with $v = p$, $k = (p-1)/2$, $\lambda = (p-3)/4$.* (These sets are non-trivial difference sets if $p \geqslant 7$, and trivial $(3, 1, 0)$ difference sets if $p = 3$).

In the general case, a simple counting argument shows that

$$k(k-1) = (v-1)\lambda. \tag{3.15}$$

Let $F : G \to \mathbb{C}$ be any complex-valued function defined on G. Its right (resp. left) *autocorrelation function* is the function $\gamma_{_F} : G \to \mathbb{C}$ (resp. ${}_{_F}\gamma : G \to \mathbb{C}$) defined by

$$\gamma_{_F}(u) := \sum_{g \in G} \overline{F(g)} F(gu) \qquad {}_{_F}\gamma(u) := \sum_{g \in G} \overline{F(g)} F(ug)\,. \tag{3.16}$$

If $\chi = \chi_{_D}$ is the characteristic function of any (a priori arbitrary) subset D of G, *i.e.*, $\chi(g) = 1$ if $g \in D$ and $\chi(g) = 0$ if $g \notin D$, then D is obviously a right (resp. left) (v, k, λ) difference set if and only if

$$\gamma_{_\chi}(u) = k \quad \text{if } u = e, \qquad \gamma_{_\chi}(u) = \lambda \quad \text{if } u \neq e \tag{3.17}$$

$$\text{resp.} \quad {}_{_\chi}\gamma(u) = k \quad \text{if } u = e, \qquad {}_{_\chi}\gamma(u) = \lambda \quad \text{if } u \neq e. \tag{3.18}$$

This yields another proof of (3.15). Indeed, for example by (3.17),

$$\begin{aligned} k^2 = \left(\sum_{g \in G} \chi(g) \right)^2 &= \sum_{g \in G} \chi(g) \sum_{h \in G} \chi(h) = \sum_{g \in G} \chi(g) \sum_{u \in G} \chi(gu) \\ &= \sum_{u \in G} \gamma_{_\chi}(u) = \gamma_{_\chi}(e) + \sum_{u \neq e} \gamma_{_\chi}(u) = k + (v-1)\lambda. \end{aligned}$$

(This was, of course, just the classical convolution product argument).

3.4. *Binary functions on finite groups with autocorrelation functions constant outside the neutral element*

By a "*binary function*" we mean complex-valued functions only taking two values, which can be arbitrary complex numbers, possibly of modulus > 1. If $\alpha \in \mathbb{C}$ and $\beta \in \mathbb{C}$ are

these values, we call such a function an $\{\alpha, \beta\}$-function. Examples of binary functions are $\{0, 1\}$-functions, $\{1, -1\}$-functions (also called ± 1 functions), etc.

Although the following result might conceivably not have appeared elsewhere in the explicit form given below, it is nevertheless probably well known (in essence) to many people in combinatorics and signal processing, at least in the case of abelian or cyclic groups.

THEOREM 3.1. *Let G be a (not necessarily abelian) finite group of order v with neutral element e, and let $F : G \to \mathbb{C}$ be any binary function with values $\{\alpha, \beta\}$, ($\alpha \neq \beta$). Then the right (resp. left) autocorrelation function of F is constant on $G\backslash\{e\}$ if and only if $F^{-1}(\alpha) := \{x \in G : F(x) = \alpha\}$ is a right (resp. left) (v, k, λ) difference set. In that case the right (resp. left) autocorrelation function has value $\gamma_0 > 0$ at $x = e$, and the same (real) value $\gamma < \gamma_0$ at all $x \in G\backslash\{e\}$, where*

$$\gamma_0 = k|\alpha|^2 + (v - k)|\beta|^2 \tag{3.19}$$

and γ is defined by any of the three (equivalent) relations

$$\gamma = \gamma_0 - (k - \lambda)|\alpha - \beta|^2 \tag{3.20}$$

$$\gamma = \lambda|\alpha|^2 + (v - 2k + \lambda)|\beta|^2 \tag{3.21}$$

$$|k\alpha + (v - k)\beta|^2 = \gamma_0 + (v - 1)\gamma, \tag{3.22}$$

so that γ satisfies the inequalities

$$-\frac{\gamma_0}{v - 1} \leqslant \gamma < \gamma_0 \,. \tag{3.23}$$

The proof is based on the following lemma, which is also useful elsewhere.

LEMMA 3.1. *Let $H : G \to \mathbb{C}$ be any complex-valued function on a (not necessarily abelian) group of order v. Put $J(x) = aH(x) + b$ with $a, b \in \mathbb{C}$. Then the right autocorrelation functions $\gamma_{_H}(u)$ and $\gamma_{_J}(u)$ satisfy*

$$\gamma_{_J}(u) = |a|^2\gamma_{_H}(u) + |b|^2 \cdot v + 2\Re\left(a\bar{b}\sum_{x\in G} H(x)\right), \tag{3.24}$$

and a similar identity holds for the left autocorrelation functions.

Proof. The proof of the lemma is straightforward. To prove the above theorem, we may just consider the right autocorrelation function, as the change of variable $x \to x^{-1}$ reduces the left case to the right one. Let $\chi := \chi_{_D}$ where $D := F^{-1}(\alpha)$. Then the theorem follows from the above lemma by straightforward calculations, upon noting that each of $\chi(x)$ and $F(x)$ can be expressed in terms of the other one from the identity $F(x) = (\alpha - \beta)\chi(x) + \beta$, since $\alpha - \beta \neq 0$. $\square$

3.5. *Binary functions on finite groups with autocorrelation functions vanishing outside the neutral element*

Let G be any finite (not necessarily abelian) group of order $v \geqslant 2$, with neutral element e. Our purpose in this section is to find all *binary* functions (and, in particular, all *unimodular binary* functions) $F : G \to \mathbb{C}$ for which the, say, right autocorrelation function $\gamma_{_F}$ satisfies $\gamma_{_F}(u) = 0$ for all $u \in G\backslash\{e\}$. Without loss we may assume that the two values of F are 1 and some $\beta \in \mathbb{C}$, ($\beta \neq 1$). As a special case of the result of section 3.4, with $\gamma = 0$, G contains a (right) (v, k, λ) difference set D such that $F(x) = 1$ if $x \in D$ and $F(x) = \beta$ if $x \in G\backslash D$. Since the case $|\beta| = 1$ (of binary *unimodular* functions) is an important special case, our purpose consists of addressing the following two problems:

A) *If $D \subset G$, ($D \neq \emptyset$), is some (right) (v, k, λ) difference set, find all those $\beta \in \mathbb{C}$ such that the binary function $F : G \to \mathbb{C}$ defined by $F(x) = 1$ if $x \in D$ and $F(x) = \beta$ if $x \in G\backslash D$ satisfies $\gamma_{_F}(u) = 0$ for all $u \in G\backslash\{e\}$.*

B) *Find all those (right) (v, k, λ) difference sets $D \subset G$ for which there exists at least one complex number β, with $|\beta| = 1$, so that the above (unimodular) function F satisfies $\gamma_{_F}(u) = 0$ for all $u \notin G\backslash\{e\}$.*

Solving Problem A): Put $\alpha = 1$ and $\gamma = 0$ in (3.21), to get

$$\lambda + (v - 2k + \lambda)|\beta|^2 + 2(k - \lambda)\Re\beta = 0. \tag{3.25}$$

Recall that $G\backslash D$ is a (right) (v', k', λ') difference set with $v' = v$, $k' = v - k$, $\lambda' = v - 2k + \lambda$. So if the coefficient $\lambda' = v - 2k + \lambda$ of $|\beta|^2$ in (3.25) is zero, then D is a "trivial" difference set (see section 3.3) for which either $k = \lambda = v$ or $k = v - 1$, $\lambda = v - 2$. In the former case the left side of (3.25) reduces to $\lambda = v$, which is impossible since $v \neq 0$. In the latter case (3.25) reduces to

$$\Re\beta = 1 - \frac{v}{2} \qquad (\text{with } v \geqslant 2) \tag{3.26}$$

and those $\beta \in \mathbb{C}$ which satisfy (3.26) constitute a vertical line. The case $k = \lambda = 0$ is excluded since we supposed $D \neq \emptyset$. The last possibility for D to be a "trivial" difference set (*i.e.*, now a singleton) is $k = 1$ and $\lambda = 0$, in which case (3.25) reduces to

$$(v - 2)|\beta|^2 + 2\Re\beta = 0. \tag{3.27}$$

If $v = 2$, then (3.27) is the same as (3.26) and the acceptable $\beta \in \mathbb{C}$ are all the $\beta = it$, ($t \in \mathbb{R}$). If $v \geqslant 3$, then the set of those $\beta \in \mathbb{C}$ which satisfy (3.27) is the circle of radius $1/(v - 2)$ with center of abscissa $-1/(v - 2)$ on the real axis.

Now if D is a *non-trivial* difference set (which easily implies $v \geqslant 7$), then in (3.25) we have $\lambda \geqslant 1$, $\lambda' \geqslant 1$ and $k - \lambda \geqslant 1$. The set of these $\beta \in \mathbb{C}$ satisfying (3.25) is then the circle of radius $\sqrt{k - \lambda}/\lambda'$ with center of abscissa $-(k - \lambda)/\lambda'$ on the real axis, ($\lambda' = v - 2k + \lambda$). Let us recapitulate the result obtained regarding problem A):

THEOREM 3.2. *Let G be a (not necessarily abelian) finite group of order $v \geqslant 2$, with neutral element e. Let $D \subset G$, ($D \neq \emptyset$), be a (right) (v, k, λ) difference set. Define a binary function $F : G \to \mathbb{C}$ by $F(x) = 1$ if $x \in D$ and $F(x) = \beta$ if $x \in G\backslash\{e\}$, where $\beta \in \mathbb{C}$ is arbitrary. Then the right autocorrelation function $\gamma_{_F}$ identically vanishes on $D\backslash\{e\}$ if and only if β is in the set $S_{_D}$ of those $\beta \in \mathbb{C}$ such that*

$$\lambda + \lambda'|\beta|^2 + 2(k - \lambda)\Re\beta = 0 \tag{3.28}$$

with $\lambda' = v - 2k + \lambda$. Also $S_{_D}$ is non-empty if and only if:

a) *Either D is the complement of a singleton, in which case $S_{_D}$ is the vertical line $\Re\beta = 1 - \frac{v}{2}$.*
b) *Or $v \geqslant 3$ and D is a singleton, in which case $S_{_D}$ is the circle of radius $1/(v-2)$ with center of abscissa $-1/(v-2)$ on the real axis.*
c) *Or D is a non-trivial difference set, in which case $S_{_D}$ is the circle of radius $\sqrt{k-\lambda}/\lambda'$ with center of abscissa $-(k-\lambda)/\lambda'$ on the real axis.*

Before proceeding to address Problem B, let us make an interesting remark: Putting $D' := G\backslash D$, then in the above theorem, whether D is a trivial or a non-trivial difference set, the set $S_{_{D'}} \subset \mathbb{C}$ is always the transform of $S_{_D}$ by the inversion whose inversion-circle is the unit circle.

Solving Problem B). This boils down to deciding whether the set $S_{_D}$ of the above theorem intersects the unit circle. So let us examine the various possibilities. If D is the complement of a singleton (with $v \geqslant 2$), the vertical line $\Re\beta = 1 - \frac{v}{2}$ intersects the unit circle if and only if $1 - \frac{v}{2} \geqslant -1$, that is, either $v = 2$ or $v = 3$ or $v = 4$. For $v = 2$ the intersection points are $\beta = \pm i$, yielding the Gauss sequences $(1, \pm i)$ of length 2 (see section 3.2). For $v = 3$ the intersection points are $\beta = \exp(\pm 2i\pi/3)$, yielding the Gauss sequences $(1, 1, j)$ and $(1, 1, j^2)$, ($j = e^{2i\pi/3}$), and their obvious modifications. For $v = 4$ the intersection point is at $\beta = -1$, yielding the ± 1 functions whose values a, b, c, d are ± 1 and satisfy $abcd = -1$. For $v = 4$ the group G can be non-cyclic. If G *is* cyclic, then we get ordinary ± 1 *sequences* of length 4 which are special cases of "Barker sequences", "Golay sequences", "Shapiro sequences", "PONS sequences", etc. If $v \geqslant 3$ and D is a singleton, again $S_{_D} \neq \emptyset$ if and only if $v = 3$ or $v = 4$, and we once again obtain the same (Gauss) sequences of length 3 and the same ± 1 functions as above.

We now come to the case when D is a non-trivial difference set, and this is the only non-trivial part (!) of this section 3.5. From assertion c) in the above theorem it easily follows by elementary calculations that $S_{_D}$ intersects the unit circle if and only if

$$v - 4(k - \lambda) \leqslant 0\,. \tag{3.29}$$

On the other hand, upon choosing $\alpha = 1$ and $\beta = -1$ in Theorem 3.1, we infer from (3.19), (3.20) and (3.23) that, for *any* (v, k, λ) difference set with $v \geqslant 3$,

$$v - 4(k - \lambda) \geqslant -\frac{v}{v-1} \tag{3.30}$$

and therefore that $v - 4(k - \lambda) \geqslant -1$, since the left side of (3.30) is an integer. Thus (3.29) can be satisfied in only two cases:

$$\textit{Case } 1. \qquad v - 4(k - \lambda) = -1 \tag{3.31}$$

$$\textit{Case } 2. \qquad v - 4(k - \lambda) = 0. \tag{3.32}$$

Before studying these two (*very important*) cases, let us first clarify some terminology. Many authoritative experts (such as [2] and [18], for example) call (v, k, λ) difference sets satisfying (3.31) "Hadamard difference sets" and those satisfying (3.32) "Menon difference sets". Other authoritative experts call those satisfying (3.32) "Hadamard difference sets" and those satisfying (3.31) "Hadamard-Paley difference sets". All these choices of names are historically justified, yet this discrepancy is unfortunate and confusing. So I humbly propose to adopt the following choices (and make everyone happy):

Definitions. *Difference sets satisfying (3.31) will be called "Hadamard-Paley difference sets". Those satisfying (3.32) will be called "Hadamard-Menon difference sets".*

Let us now study these two types of difference sets and their incidences on our Problem B). Elementary arithmetical calculations show that (assuming $k \leqslant v/2$ *without loss*) the diophantine identities (3.15) and (3.31) hold simultaneously if and only if the parameters (v, k, λ) of the (Hadamard-Paley) difference set have the form

$$v = 4n - 1, \qquad k = 2n - 1, \qquad \lambda = n - 1 \tag{3.33}$$

and that the diophantine identities (3.15) and (3.32) hold simultaneously if and only if the parameters (v, k, λ) of the (Hadamard-Menon) difference set have the form

$$v = 4N^2, \qquad k = 2N^2 - N, \qquad \lambda = N^2 - N\,. \tag{3.34}$$

Now some elementary calculations show that if D is a (right, say) Hadamard-Paley difference set as defined by (3.33), then the only *unimodular* solutions $\beta \in \mathbb{C}$ of our above Problem B) are

$$\beta = \exp\left(i \arccos \frac{1-v}{1+v}\right) = \exp\left(i \arccos \frac{1-2n}{1+2n}\right) \tag{3.35}$$

and, of course, the complex conjugate of this number. Also elementary calculations show that if D is a (right, say) Hadamard-Menon difference set as defined by (3.34), then the only unimodular solution $\beta \in \mathbb{C}$ of our above Problem B) is $\beta = -1$.

Also observe that all the solutions of Problem B) in the case of *trivial* difference sets are special cases of the above discussion (for Hadamard-Paley and Hadamard-Menon difference sets), *except* for the Gaussian sequences $(1, \pm i)$ of length 2. Let us now recapitulate:

THEOREM 3.3. *Let G be any finite (not necessarily abelian) finite group of order $v \geqslant 2$, with neutral element e. Let $D \subset G$, ($D \neq \emptyset$), be a (right) (v, k, λ) difference set (satisfying $k \leqslant v/2$ without loss). Define a binary unimodular function $F : G \to \mathbb{C}$ by $F(x) = 1$ if $x \in D$ and $F(x) = \beta$ if $x \in G \backslash D$ where $\beta \in \mathbb{C}$ and $|\beta| = 1$. Then the (right) autocorrelation function γ_F satisfies $\gamma_F(u) = 0$ for all $u \in G \backslash \{e\}$ if and only if we are in one of these three situations:*

a) *$v = 2$ and F is one of the (Gauss) sequences $(1, i)$ and $(1, -i)$.*
b) *D is a Hadamard-Paley difference set given by (3.33) and β is defined by (3.35) or is the complex conjugate of that number.*
c) *D is a Hadamard-Menon difference set given by (3.34), and $\beta = -1$.*

3.6. *Binary bi-unimodular sequences*

The case of binary bi-unimodular *sequences* is just the special case of the discussions and results of section 3.5 when the group G is *cyclic*. The extension of the result (3.12) of Björck [6], promised in section 3.2, is the special case of Theorem 3.3 when G is cyclic.

This now leads us to the crucial problem of the *existence and determination* of such Hadamard-Paley and Hadamard-Menon difference sets. Even in the case of cyclic groups these existence problems are far from being entirely solved. Several classes of *cyclic* Hadamard-Paley difference sets are known. See *for example* [2] and [18]. As for *cyclic* Hadamard-Menon difference sets, no example is known (except, of course, for $v = 4$), and a long-standing conjecture (due to Ryser [33]) is that none exists if $v > 4$. This is equivalent to the famous "*Hadamard circulant conjecture*" which states that no circulant matrix of order v with entries ± 1 can be a Hadamard matrix *unless* $v = 4$.

The interpretation of binary bi-unimodular sequences in terms of polynomials of degree $v - 1$ (with binary unimodular coefficient sequences), which are perfectly flat on the group Γ_v of v-th roots of 1, is by now obvious. We will come back to these in some depth in [39].

3.7. *Sup-norms of bi-unimodular polynomials on the unit circle*

A "*bi-unimodular polynomial*" is, of course, the associated polynomial

$$P(X) = \sum_{k=0}^{L-1} a_k X^k \tag{3.36}$$

of a bi-unimodular sequence $(a_0, a_1, \ldots, a_{L-1})$.

A useful theorem of Landau [21] says that if $Q(X) = \sum_{k=0}^{L-1} c_k X^k \in \mathbb{C}[X]$ is *any* polynomial with complex coefficients, then its maximum modulus $\|Q\|_\infty$ on the whole unit circle

is majorized in terms of its maximum modulus on the group Γ_L of L-the roots of 1 as follows:

$$\|Q\|_\infty \leqslant C \cdot \left(\max_{g \in \Gamma_L} |Q(g)| \right) \cdot \log L \tag{3.37}$$

where C is some absolute constant, and the $\log L$ cannot be replaced (in the general case) by a smaller factor. This can be proved by Lagrange interpolation, and also otherwise.

Since our bi-unimodular polynomial (3.36) satisfies $|P(g)| = \sqrt{L}$ for all $g \in \Gamma_L$, (3.37) implies that

$$\|P\|_\infty \leqslant C\sqrt{L} \log L. \tag{3.38}$$

After Björck [6] gave, at the 1989 ASI on "*Recent Advances in Fourier Analysis*", his lecture on (what we subsequently called the) bi-unimodular sequences, one of the participants (*perhaps* H. S. Shapiro or D. J. Newman, if my memory is correct) conjectured that, for *bi-unimodular* polynomials $P(X)$, (3.38) could be improved to:

$$\|P\|_\infty \leqslant C_1\sqrt{L} \qquad (C_1 = \text{some absolute constant}), \tag{3.39}$$

i.e., $P(X)$ has *bounded crest factor* (see beginning of part 3). His "evidence" for the conjecture (3.39) was twofold:

A) On one hand the conjecture (3.39), in the special case of the Gauss sequences

$$a_k = \omega^{ak^2+bk} \tag{3.40}$$

(L odd, $\omega = e^{2i\pi/L}$, a relatively prime to L), had been around for many decades, perhaps even before the 1920's and, incidentally, is *still open* in 2000–2001.

B) On the other hand (3.39) was known to be true for two particular classes of bi-unimodular sequences/polynomials, namely the Gauss sequences (3.40) with the choices $a = (L \pm 1)/2$ (see Littlewood [22]) and the bi-unimodular polynomials (of length $L = M^2$) studied in 1977 by Byrnes [10]:

$$P(X) = \sum_{h=0}^{M-1} \sum_{r=0}^{M-1} e^{2hri\pi/M} X^{Mh+r}. \tag{3.41}$$

I did some thinking on the conjecture (3.39) from time to time, until I saw how to *disprove* it in June 2000:

THEOREM 3.4. *If p is a prime $\equiv -1 (\mathrm{mod} 4)$ and a_k is the binary bi-unimodular sequence of length p (originally introduced by Björck) defined by*

$$a_k = \exp\left(i \arccos \frac{1-p}{1+p} \right)$$

if k is a quadratic non-residue (mod p) and $a_k = 1$ otherwise, ($0 \leqslant k \leqslant p-1$), then the associated (bi-unimodular) polynomial satisfies, for all sufficiently large p,

$$\|P\|_\infty > \frac{2}{\pi}\sqrt{p}\log\log p - \frac{1}{2} \tag{3.42}$$

and, if $\epsilon > 0$ is fixed, then for infinitely many such primes p

$$\|P\|_\infty > \left(\frac{2}{\pi}\cdot e^\gamma - \epsilon\right)\sqrt{p}\log\log p \tag{3.43}$$

where γ is Euler's constant.

These results are immediate consequences of Montgomery's 1980 work [28] on the Fekete polynomials $\sum_{k=0}^{p-1}\chi\left(\frac{k}{p}\right)X^k$, where $\chi\left(\frac{k}{p}\right)$ is Legendre's symbol. It is unforgivable that I have not seen earlier this straightforward disproof of conjecture (3.39), since Montgomery had given me a copy of his paper [28] as early as 1988.

This disproof, in turn, gives rise to interesting results and problems on binary bi-unimodular sequences, which I hope to discuss in [39].

4. The drunkard, the bar owner and the police

Sorry, dear reader, I must abruptly stop my write-up at this point. I have a deadline to respect and I had underestimated the time and space needed to just *touch upon* some most interesting topics I had in mind when writing the introduction. Below is a list of *some* of these topics, and I am also omitting some of them, with the intention of including most of them in the extended version [39] of this paper (which, in a way, I have barely started to write).

4.1. *Some topics to be found in the extended version [39]*

- The genesis of the *analytic* theory of flat polynomials from nineteenth century Fourier analysis. The origins from the absolute convergence of trigonometric and Fourier series: Dirichlet, Fejér, Bernstein and many others up to the present days.
- The various Hardy-Littlewood unimodular polynomials with bounded crest factors. Their connections with number theory and Fourier analysis.
- Bounded crest factor properties of Gauss and Byrnes polynomials and open problems. The van der Corput methods and their discrete analogs (Kuzmin, Landau). D.J. Newman's method.
- D. J. Newman and H. S. Shapiro in New York in the late 1940's. The making of the *Shapiro polynomials* at M.I.T. (1950–1951). Raphaël Salem and the Fekete polynomials. Infra-red spectrometry and the Golay sequences. History of Shapiro polynomials and Golay sequences: Rediscoveries, truths, lies, counter-truths and folklore galore.
- Barker sequences and their generalizations. The Hadamard Circulant Conjecture, from Reiser to B. Schmidt via R. Turyn.

- D.J. Newman's conjecture and Barker sequences. Recent (published and unpublished) research.
- Norms of unimodular polynomials. History of Erdös-Newman conjectures. Littlewood's work and his various conjectures and "*counter-conjectures*". The strange story of the Byrnes pseudo-polynomials, Körner's "true-false" theorem, Kahane's ultra-flat polynomials and further on-going research.
- Littlewood's conjecture on L^p-norms of exponential sums: Exciting research from Paul Cohen (late 1950's) to Ivo Klemes (early 2000's).

4.2. *Time to stop*

The above list is incomplete. It is pointless to try to remember all the topics that I had *intended* to put in this paper, as I have to stop anyway. I deliberately *refrained from giving any references* in the above section 4.1, as there is just too much exciting work (whether old, recent, unpublished or on-going) to quote. It is better to do a rather thorough job in [39] than a sloppy job here.

I find myself in the same situation as a drunkard sitting late at night in a bar with a (nearly) full bottle of whiskey he intends to drink, who is suddenly told by the bar owner that it's now closing time and that the police are about to arrive any second and enforce the regulations. The drunkard has to leave the bar, and I have to stop here. By the time this (quasi-aborted) paper appears, I hope the extended version [39] will be available for whoever is interested.

References

[1] J. Arsac. *Fourier Transforms and the Theory of Distributions*. Prentice Hall, 1966.

[2] L. D. Baumert. *Cyclic Difference Sets*. Lecture Notes in Mathematics 182, Springer-Verlag, 1971.

[3] S. N. Bernstein. *On the order of the best approximation of continuous functions by polynomials of given degree*. Memoirs of the Royal Academy of Belgium (2), 4, pages 1–103, 1912. (French)

[4] S. N. Bernstein. *Lectures on extremal properties and the best approximation of analytic functions of a real variable* Gauthiers-Villars, 1926. (Reprinted as the first part of "*Approximation*" by S. N. Bernstein and C. de La Vallée Poussin, Chelsea, New York, 1970). (French).

[5] G. Björck. Functions of modulus one on $\mathbb{Z}_p$ whose Fourier transforms have constant modulus. *Proceedings of A. Haar Memorial Conference*, 1985, Colloq. Math. Soc. János Bolyai 49, pages 193–197, 1985.

[6] G. Björck. Functions of modulus 1 on $\mathbb{Z}_n$ whose Fourier transforms have constant modulus, and "cyclic n-roots". *Proceedings of the 1989 NATO Advanced Study Institute on "Recent Advances in Fourier Analysis and Its Applications*", J. S. Byrnes & J. L. Byrnes, ed., pages 131–140, 1990.

[7] G. Björck & B. Saffari. *New classes of finite unimodular sequences with unimodular Fourier transforms. Circulant Hadamard matrices with complex entries*. C. R. Acad. Sci. Paris, Ser. I Math. 320, No. 3, pages 319–324, 1995.

[8] R. P. Boas, Jr. Inequalities for the derivatives of polynomials, *Math. Mag.* 42, pages 165–174, 1969.

[9] P. Borwein. Home page, with (fairly) complete list of publications, on Web site of Math. Dept. , Simon Fraser University. Updated 2000-2001. (*c.f.* Ref. [14]).

[10] J. S. Byrnes. On polynomials with coefficients of modulus one. Bull. London Math. Soc. 9, pages 171–176, 1977.

[11] C. De La Vallée Poussin. Sur la convergence des formules d'interpolation entre ordonnées équidistantes, Acad. Roy. Belg. Bull. Cl. Sci (5), pages 319–410, 1908. (French)

[12] C. De La Vallée Poussin. *Lectures on approximation of functions of a real variable*, 1919. (Reprinted as the second part of "Approximation" by S. N. Bernstein and C. de La Vallée Poussin, Chelsea, New York, 1970). (French)

[13] P. Enflo. Oral question about a uniqueness property of Gauss sequences of prime length, Stockholm University, 1983.

[14] T. Erdelyi. Home, page with (fairly) complete list of publications, on Web site of Math. Dept., Texas A&M University. Updated 2000–2001. (*c.f.* Ref. [9]).

[15] L. Fejér. *Über konjugierte trigonometrische Reihen*, J. Reine Angew. Math. 144, pages 48–56, 1914.

[16] M. Fekete. *Über einen Satz des Herrn Serge Bernstein*, J. Reine Angew. Math. 146, pages 88–94, 1916.

[17] C. Frappier. Some extremal problems for polynomials and entire functions of exponential type. Ph. D. thesis, University of Montreal, 1983. (French).

[18] S. W. Golomb. Cyclic Hadamard difference sets. *SETA'*98 (*Proceedings of December* 1998 *Conference on "Sequences and their applications", Singapore*). World Scientific, 2000.

[19] U. Haagerup. Personal letter (in Danish) to G. Björck, with copy to B. Saffari, 1995.

[20] E. Landau. Personal letter to S. N. Bernstein, 1912.

[21] E. Landau. Bemerkungen zu einer Arbeit von Herrn Carlemann. Math. Zeit. Bd 5, 1919.

[22] J. E. Littlewood. *Some Problems in Real and Complex Analysis*, Heath Mathematical Monographs, Heath & Co., 1968.

[23] G. G. Lorentz. *Approximation of Functions*. Holt, Rinehart & Winston, 1966.

[24] L. Lovász. Private communication to G. Björck, 1983.

[25] A. A. Markov. On a question posed by D. I. Mendeleiev. Bulletin of the Academy of Sciences of St. Petersburg 62, pages 1–24, 1889. (Reprinted in "Selected Works", Izdat. Akad. SSSR, Moscow, pages 51–75, 1948. (Russian)

[26] Z. A. Melzak. *Companion to Concrete Mathematics*, vol II, Wiley, 1976.

[27] D. I. Mendeleiev. Study of aqueous dissolutions, based on changes of their specific weights. St. Petersburg, 1887. (Russian).

[28] H. L. Montgomery. An exponential polynomial formed with the Legendre symbol. Acta Arithmetica, vol 38, pages 375–380, 1980.

[29] R. E. A. C. Paley. On Orthogonal Matrices, J. Math. and Phys. 12, pages 311-320, 1933.

[30] S. K. Pichorides & B. Saffari. A proof by interpolation of a theorem of Szegö. Private manuscript, 1980. (presented in Ref. [36], 1980–1981).

[31] Q. I. Rahman. Applications of Functional Analysis to Extremal Problems for Polynomials. Presses. Univ. Montréal, 1968.

[32] Q. I. Rahman & G. Schmeisser. Les inégalités de Markov et de Bernstein. Presses Univ. Montréal, 1983.

[33] H. J. Ryser. Combinatorial Mathematics. Carus Math. Monographs 14, Wiley, 1963.

[34] F. Riesz. On trigonometric polynomials. C. R. Acad. Sci. Paris, vol 158, pages 1657–1661, 1914. (French)

[35] M. Riesz. Interpolation formula for the derivative of a trigonometric polynomial. C. R. Acad. Sci. Paris vol. 158, pp. 1152–1154, 1914. (French)

[36] B. Saffari. An interpolation proof of Szegö's inequality by S. K. Pichorides and myself. In "*Extremal problems on trigonometric polynomials*", Postgraduate Course at Univ. of Geneva, Switzerland, 1980–1981.

[37] B. Saffari. New unimodular polynomials with vanishing periodic autocorrelations. Research report, Prometheus Inc., Newport, RI, 1990.

[38] B. Saffari. Perfect flatness of polynomials with coefficients ± 1 on groups of roots of unity. Private manuscript, 1998.

[39] B. Saffari, Twentieth century extremal problems on polynomials and exponential sums, extended version of this paper, (in preparation).

[40] G. Schaake & J. G. van der Corput. Ungleichungen für Polynome und trigonometrische Polynome. *Compositio Math.* vol 2, pages 321–361, 1935.

[41] G. Schaake & J. G. van der Corput. Berichtigung zu: Ungleichungen für Polynome und trigonometrische Polynome. Compositio Math. vol 3, page 128, 1936.

[42] G. Szegö. Über einen Satz des Herrn Serge Bernstein. Schriften der Königsberger Gelehrten Gesellschaft, pages 59–70, 1928.

[43] G. Szegö & A. Zygmund. On certain mean values of Polynomials, J. d'Analyse Math. vol 3, pages 225–244, 1953–1954.

[44] A. Zygmund. Trigonometric Series. Cambridge at the University Press, 1959.